PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/705 | A2

(11) International Publication Number: **WO 95/11974**

(43) International Publication Date: 4 May 1995 (04.05.95)

(21) International Application Number: PCT/US94/11897

(22) International Filing Date: 19 October 1994 (19.10.94)

(30) Priority Data:

08/141,500	22 October 1993 (22.10.93)	US
08/143,215	26 October 1993 (26.10.93)	US

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

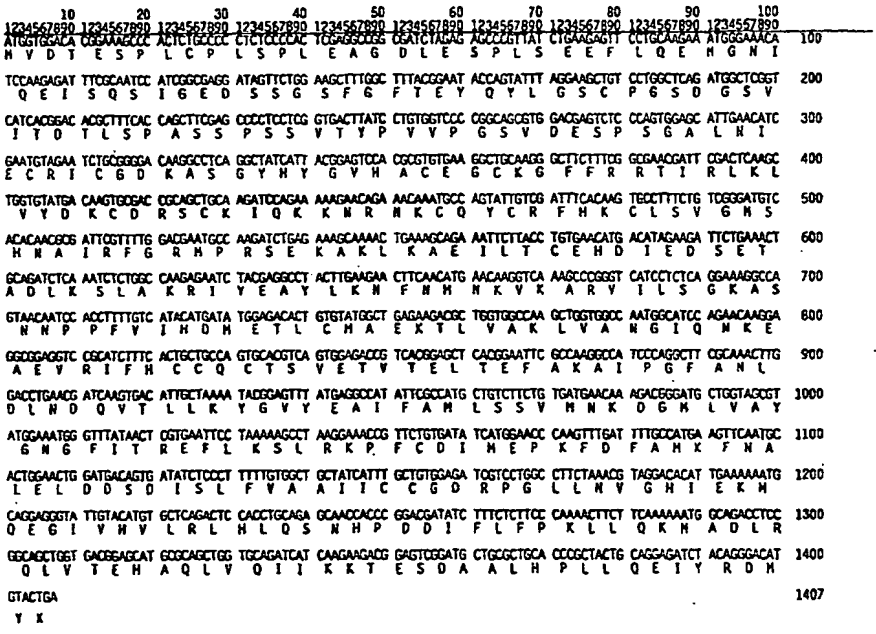
(71) Applicant: LIGAND PHARMACEUTICALS, INC. [US/US]; 9393 Towne Center Drive, San Diego, CA 92121 (US).

(72) Inventor: MUKHERJEE, Ranjan; 11341 Avenida De Los Lobos, San Diego, CA 92127 (US).

(74) Agents: CHEN, Anthony, C. et al.; Lyon & Lyon, First Interstate World Center, Suite 4700, 633 West Fifth Street, Los Angeles, CA 90071-2066 (US).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR



(57) Abstract

A human peroxisome proliferation activated receptor gene is purified from the environment in which it naturally occurs, and preferably provided within an expression vector.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

DESCRIPTION

Human Peroxisome Proliferator Activated Receptor

Cross Reference to Related Application

This application is a continuation-in-part of Application Docket No. 202/041, titled "Human Peroxisome Proliferator Activated Receptor," filed October 22, 1993, by Mukherjee, the disclosure of which is incorporated herein by reference.

Field of the Invention

This invention relates to the cloning and uses of a human peroxisome proliferator activated receptor.

Background of the Invention

A peroxisome proliferator is an agent that induces peroxisomal proliferation. Peroxisome proliferators are a diverse group of chemicals which include unsaturated fatty acids, hypolipidemic drugs, herbicides, leukotriene antagonists, and plasticizers (for a review, see Green, S., 43 Biochem. Pharmacol. 393-400, 1992). Hypolipidemic drugs such as clofibrates have been found to lower triglycerides and cholesterol levels in plasma and to be beneficial in the prevention of ischaemic heart disease in individuals with elevated levels of cholesterol (Havel, R.J. and Kane, J.P., 13 Ann. Rev. Pharmac. 287-308, 1973). Therapeutic use of such drugs, however, is questioned because clofibrates are carcinogens in rats.

Peroxisome proliferator activated receptor (PPAR) is a member of the steroid receptor family. It is activated by peroxisome proliferators. Issemann and Green, 347 Nature 645, 1990, cloned a mouse peroxisome proliferator activated receptor (mPPAR) gene from a mouse liver complementary DNA (cDNA) library. Göttlicher et al., 89 Proc. Nat. Acad. Sci. USA 4653-4657, 1992, cloned a rat peroxisome proliferator activated receptor (rPPAR) gene from a rat liver cDNA library. PPARs from mouse and rat share 97% homology in amino acid sequence and a

particularly well-conserved putative ligand-binding domain. Three members of the Xenopus nuclear hormone receptor superfamily have also been found to be structurally and functionally related to the mPPAR (Dreyer et al., 68 Cell 879-887, 1992).

Schmidt et al., 6 Molecular Endocrinology 1634-1641, 1992, cloned a steroid hormone receptor gene, NUC1, from a human osteosarcoma cell cDNA library. The homology between amino acid sequence of NUC1 and that of the mouse PPAR is only 62%. Thus, although it is clear that NUC1 is a member of the PPAR receptor group, it remains to be determined whether NUC1 is the human homolog of the mouse PPAR or a new member of the PPAR family.

Sher et al., 32 Biochemistry 5598-5604, 1993, cloned a human PPAR gene from a human liver cDNA library. This clone has 85% nucleotide sequence homology and 91% amino acid sequence homology with the mPPAR clone.

Summary of the Invention

The present invention relates to the cloning of a human PPAR gene, hPPAR1. The protein encoded by hPPAR1 has 92% homology with the mouse PPAR. It is different from the human PPAR cloned by Sher et al., supra, at two locations in the amino acid sequence, i.e., amino acids 268 and 296.

The hPPAR1 clone can be used for the expression of large amounts of hPPAR1. This human PPAR clone is also useful for screening compounds for improved pharmacological profiles for the treatment of hyperlipidemia with higher potency, efficacy, and fewer side effects. Specifically, the human PPAR clone can be used to screen for compounds active as primary endogenous inducers of the human PPAR. In addition, it is useful for establishing the tissue specific expression pattern of human PPAR. For example, a Northern blot can be used to reveal tissue specific expression of the gene to aid treatment of diseases such as atherosclerosis.

Thus, in a first aspect, the invention features a purified nucleic acid encoding a human PPAR with the nucleotide base sequence shown in Figure 1, and given as SEQ ID NO. 1. By purified nucleic acid is meant that the nucleic acid is separated from its natural environment and from other nucleic acids.

In a second aspect, the present invention features a vector containing the human PPAR gene. This vector may be used for multiplication of the human PPAR gene or expression of the human PPAR gene.

In a preferred embodiment, the vector is an expression vector. In one example, the expression vector is used to make a recombinant human PPAR nucleic acid, which can be used as a specific probe for DNA or RNA complementary to the human PPAR sequence. In another example, the expression vector is used to express human recombinant PPAR protein.

By vector is meant a plasmid or viral DNA molecule into which either a cDNA or a genomic DNA sequence is inserted.

By expression vector is meant a vector that directs protein synthesis from a promoter. In a preferred embodiment, either vector pBacPAK8 (Clontech) or vector pBacPAK9 (Clontech) is used to express the human PPAR in insect cells. In another preferred embodiment, vector pYES2 (Invitrogen) is used to express the human PPAR in yeast cells. In yet another preferred embodiment, pBKCMV (Stratagene) is used to express the human PPAR in mammalian cells.

By recombinant human PPAR is meant a non-naturally expressed human PPAR.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Description of the Preferred Embodiments

Drawings

Figure 1 is the nucleotide and amino acid sequence of hPPAR1; and

Figure 2 is a comparison of the amino acid sequences of hPPAR1 and the mouse PPAR.

What follows is an example of the cloning of a human PPAR. Those of ordinary skill in the art will recognize that equivalent procedures can be readily used to isolate human PPAR from cDNA libraries or genomic libraries of other tissues than that exemplified below, namely the liver.

In general, the cloning of the human PPAR involved probing a human liver cell cDNA library with a labeled EcoRI-BglII fragment (nucleotides 450-909) of the rat PPAR (459 bases). The sequence of the probe is shown in Göttlicher et al. supra.

The recipes for buffers, mediums, and solutions in the following examples are given in J. Sambrook, E. F. Fritsch, and T. Maniatis, Molecular Cloning: A Laboratory Manual, 2 Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989.

Example 1: Cloning of a human PPAR

A human PPAR subtype, hPPAR1, was cloned from a human liver 5'-stretch cDNA library (Clontech #HL1115a) in lambda gt10 phages. C600-Hfl coli (Clontech) was grown overnight in LB broth supplemented with 0.2% maltose. A required amount of phage (corresponding to 2 million plaques) was mixed with 200 microliters of 10 mM $MgCl_2$/10 mM $CaCl_2$ and 1.5 milliliters of the overnight C600-Hfl coli and incubated at 37°C for 30 minutes. Soft LB agarose was added at 48°C, mixed and poured onto prewarmed 22x22 cm rectangular LB agar plates and incubated overnight at 37°C.

Plaque lifts were performed by chilling the plates at 4°C to harden the top agarose and prevent it from peeling,

marking a nylon or nitrocellulose filter on the surface contacting the plaques, laying the filter on the surface without trapped air bubbles, and leaving it for about a minute. A number of asymmetric dots were inserted with Indian ink with a syringe and needle so that the ink soaked into the agar. The sheets were then peeled gently away, and laid plaque side up on two sheets of Whattman 3MM soaked in denaturing solution, and left for about 2 minutes. The sheets were then peeled away and immersed in a standard neutralizing solution for 5 minutes, immersed in 5X SSC, air dried, and baked at 80°C under vacuum, for 2 hours.

The filters were prehybridized in 40% formamide, 5X SSC, 0.1 % SDS, 1X Denhardt, and 100 ng/ml denatured salmon sperm DNA at 37°-42°C for 1 hour. Labeled DNA probe (1 million cpm/ml) was denatured by heating at 100°C for 10 minutes, chilled, and then added to the prehybridization solution, and hybridized at 37°-42°C overnight. The filters were washed in 2X SSC and, 0.1% SDS at 42°C or higher temperature.

Positive plaques were identified and purified by rescreening two more times. The probe was labeled by nick-translation.

Phage stocks were made by isolating and removing a well separated plaque with the narrow end of an autoclaved Pasteur pipette, immersing it in 1 ml of standard SM buffer, and adding a drop of chloroform. This was left for a few hours at room temperature (20°C-24°C) or overnight at 4°C, vortexed, and centrifuged.

The cDNA insert was amplified by polymerase chain reactions (PCR). 20 microliters of phage stock was used in 100 microliters of standard PCR reaction buffer, by adding all components except Polymerase. This mixture was heated to 99°C, and Vent DNA polymerase (Biolabs) was added to start the PCR cycles. The PCR conditions were 95°C 1 minute, 72°C 1 minute, 72°C 3 minutes (1 minute per

kilobase) for 30 cycles, 72°C 5 minutes, and kept at 4°C till further utilized.

The applicant isolated a clone from the cDNA library using an EcoR1-BglII fragment (nucleotides 450-909) of the rat PPAR (459 bases) as a probe and the hybridization conditions provided above. This clone was purified and its sequence defined. This sequence is shown in Figure 1, and as SEQ. ID. NO. 1. Figure 2 is a comparison of mPPAR and hPPAR1 amino acid sequences with those amino acids having identity between the two sequences enclosed in blocks.

Example 2: Northern blot analysis

A human multiple tissue Northern blot was purchased from Clontech. Screening was done following the manufacturer's protocol. The blot was prehybridized in 5X SSPE, 10X Denhardt's solution, 100μg/ml of freshly denatured salmon sperm DNA, 50% formamide and 2% SDS for 3 hours at 42°C. DNA from the EcoR1 site at position 1025 of the coding region to the end of the cloned gene was used as probe (see Figure 1). This DNA was labeled by random priming, boiled and added at a concentration of 1 million cpm/ml of prehybridization solution. Hybridization was carried out for 13 hours at 42°C. The blot was then washed in 2X SSC, 0.05% SDS at room temperature followed by two washes in 0.1X SSC, 0.1% SDS at 50°C and exposed to X-ray film.

A specific band of about 10 kilobase was observed in all tissues except the brain. Maximal expression was observed in skeletal muscle, followed by heart, placenta, pancreas, liver, kidney, and lung. The expression of hPPAR1 gene is therefore observed in tissues known to express PPARs in other species.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: LIGAND PHARMACEUTICALS, INC.
(B) STREET: 9393 Towne Centre Drive
(C) CITY: San Diego
(D) STATE: California
(E) COUNTRY: United States of America
(F) POSTAL CODE (ZIP): 92121
(G) TELEPHONE: (619) 535-3900
(H) TELEFAX: (619) 535-3906

(ii) TITLE OF INVENTION: HUMAN PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR

(iii) NUMBER OF SEQUENCES: 3

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
(B) COMPUTER: IBM compatible
(C) OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
(D) SOFTWARE: WordPerfect (Version 5.1)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: To Be Assigned

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/141,500
(B) FILING DATE: 22-OCT-1993

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/143,215
(B) FILING DATE: 26-OCT-1993

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1407 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION : SEQ ID NO: 1:

```
ATG GTG GAC ACG GAA AGC CCA CTC TGC CCC CTC TCC CCA     39
Met Val Asp Thr Glu Ser Pro Leu Cys Pro Leu Ser Pro
                5                  10

CTC GAG GCC GGC GAT CTA GAG AGC CCG TTA TCT GAA GAG     78
Leu Glu Ala Gly Asp Leu Glu Ser Pro Leu Ser Glu Glu
     15                 20                  25

TTC CTG CAA GAA ATG GGA AAC ATC CAA GAG ATT TCG CAA    117
Phe Leu Gln Glu Met Gly Asn Ile Gln Glu Ile Ser Gln
            30                  35

TCC ATC GGC GAG GAT AGT TCT GGA AGC TTT GGC TTT ACG    156
Ser Ile Gly Glu Asp Ser Ser Gly Ser Phe Gly Phe Thr
 40                 45                  50

GAA TAC CAG TAT TTA GGA AGC TGT CCT GGC TCA GAT GGC    195
Glu Tyr Gln Tyr Leu Gly Ser Cys Pro Gly Ser Asp Gly
        55                  60                  65

TCG GTC ATC ACG GAC ACG CTT TCA CCA GCT TCG AGC CCC    234
Ser Val Ile Thr Asp Thr Leu Ser Pro Ala Ser Ser Pro
                70                  75

TCC TCG GTG ACT TAT CCT GTG GTC CCC GGC AGC GTG GAC    273
Ser Ser Val Thr Tyr Pro Val Val Pro Gly Ser Val Asp
     80                 85                  90

GAG TCT CCC AGT GGA GCA TTG AAC ATC GAA TGT AGA ATC    312
Glu Ser Pro Ser Gly Ala Leu Asn Ile Glu Cys Arg Ile
            95                 100

TGC GGG GAC AAG GCC TCA GGC TAT CAT TAC GGA GTC CAC    351
Cys Gly Asp Lys Ala Ser Gly Tyr His Tyr Gly Val His
105                 110                 115

GCG TGT GAA GGC TGC AAG GGC TTC TTT CGG CGA ACG ATT    390
Ala Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Thr Ile
        120                 125                 130

CGA CTC AAG CTG GTG TAT GAC AAG TGC GAC CGC AGC TGC    429
Arg Leu Lys Leu Val Tyr Asp Lys Cys Asp Arg Ser Cys
                135                 140

AAG ATC CAG AAA AAG AAC AGT TTC AAA TGC CAG TAT TGT    468
Lys Ile Gln Lys Lys Asn Arg Asn Lys Cys Gln Tyr Cys
    145                 150                 155

CGA TTT CAC AAG TGC CTT TCT GTC GGG ATG TCA CAC AAC    507
Arg Phe His Lys Cys Leu Ser Val Gly Met Ser His Asn
            160                 165
```

```
GCG ATT CGT TTT GGA CGA ATG CCA AGA TCT GAG AAA GCA   546
Ala Ile Arg Phe Gly Arg Met Pro Arg Ser Glu Lys Ala
170                 175                 180

AAA CTG AAA GCA GAA ATT CTT ACC TGT GAA CAT GAC ATA   585
Lys Leu Lys Ala Glu Ile Leu Thr Cys Glu His Asp Ile
        185                 190                 195

GAA GAT TCT GAA ACT GCA GAT CTC AAA TCT CTG GCC AAG   624
Glu Asp Ser Glu Thr Ala Asp Leu Lys Ser Leu Ala Lys
                200                 205

AGA ATC TAC GAG GCC TAC TTG AAG AAC TTC AAC ATG AAC   663
Arg Ile Tyr Glu Ala Tyr Leu Lys Asn Phe Asn Met Asn
    210                 215                 220

AAG GTC AAA GCC CGG GTC ATC CTC TCA GGA AAG GCC AGT   702
Lys Val Lys Ala Arg Val Ile Leu Ser Gly Lys Ala Ser
225                 230

AAC AAT CCA CCT TTT GTC ATA CAT GAT ATG GAG ACA CTG   741
Asn Asn Pro Pro Phe Val Ile His Asp Met Glu Thr Leu
235                 240                 245

TGT ATG GCT GAG AAG ACG CTG GTG GCC AAG CTG GTG GCC   780
Cys Met Ala Glu Lys Thr Leu Val Ala Lys Leu Val Ala
        250                 255                 260

AAT GGC ATC CAG AAC AAG GAG GCG GAG GTC CGC ATC TTT   819
Asn Gly Ile Gln Asn Lys Glu Ala Glu Val Arg Ile Phe
                265                 270

CAC TCG TGC CAG TGC ACG TCA GTG GTG ACC GTC ACG GAG   858
His Cys Cys Gln Cys Thr Ser Val Glu Thr Val Thr Glu
    275                 280                 285

CTC ACG GAA TTC GCC AAG GCC ATC CCA GGC TTC GCA AAC   897
Leu Thr Glu Phe Ala Lys Ala Ile Pro Gly Phe Ala Asn
290                 295

TTG GAC CTG AAC GAT CAA GTG ACA TTG CTA AAA TAC GGA   936
Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly
300                 305                 310

GTT TAT GAG GCC ATA TTC GCC ATG CTG TCT TCT GTG ATG   975
Val Tyr Glu Ala Ile Phe Ala Met Leu Ser Ser Val Met
        315                 320                 325

AAC AAA GAC GGG ATG CTG GTA GCG TAT GGA AAT GGG TTT  1014
Asn Lys Asp Gly Met Leu Val Ala Tyr Gly Asn Gly Phe
                330                 335

ATA ACT CGT GAA TTC CTA AAA AGC CTA AGG AAA CCG TTC  1053
Ile Thr Arg Glu Phe Leu Lys Ser Leu Arg Lys Pro Phe
    340                 345                 350
```

```
TGT GAT ATC ATG GAA CCC AAG TTT GAT TTT GCC ATG AAG  1092
Cys Asp Ile Met Glu Pro Lys Phe Asp Phe Ala Met Lys
            355                 360

TTC AAT GCA CTG GAA CTG GAT GAC AGT GAT ATC TCC CTT  1131
Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Ile Ser Leu
365                 370                 375

TTT GTG GCT GCT ATC ATT TGC TGT GGA GAT CGT CCT GGC  1170
Phe Val Ala Ala Ile Ile Cys Cys Gly Asp Arg Pro Gly
        380                 385                 390

CTT CTA AAC GTA GGA CAC ATT GAA AAA ATG CAG GAG GGT  1209
Leu Leu Asn Val Gly His Ile Glu Lys Met Gln Glu Gly
                395                 400

ATT GTA CAT GTG CTC AGA CTC CAC CTG CAG AGC AAC CAC  1248
Ile Val His Val Leu Arg Leu His Leu Gln Ser Asn His
    405                 410                 415

CCG GAC GAT ATC TTT CTC TTC CCA AAA CTT CTT CAA AAA  1287
Pro Asp Asp Ile Phe Leu Phe Pro Lys Leu Leu Gln Lys
            420                 425

ATG GCA GAC CTC CGG CAG CTG GTG ACG GAG CAT GCG CAG  1326
Met Ala Asp Leu Arg Gln Leu Val Thr Glu His Ala Gln
430                 435                 440

CTG GTG CAG ATC ATC AAG AAG ACG GAG TCG GAT CGT GCG  1365
Leu Val Gln Ile Ile Lys Lys Thr Glu Ser Asp Ala Ala
        445                 450                 455

CTG CAC CCG CTA CTG CAG GAG ATC TAC AGG GAC ATG TAC  1404
Leu His Pro Leu Leu Gln Glu Ile Tyr Arg Asp Met Tyr
                460                 465

TGA                                                  1407
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 468 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION : SEQ ID NO: 2

```
Met Val Asp Thr Glu Ser Pro Leu Cys Pro Leu Ser Pro
                5                  10

Leu Glu Ala Gly Asp Leu Glu Ser Pro Leu Ser Glu Glu
    15                  20                  25
```

Phe Leu Gln Glu Met Gly Asn Ile Gln Glu Ile Ser Gln
30 35

Ser Ile Gly Glu Asp Ser Ser Gly Ser Phe Gly Phe Thr
40 45 50

Glu Tyr Gln Tyr Leu Gly Ser Cys Pro Gly Ser Asp Gly
55 60 65

Ser Val Ile Thr Asp Thr Leu Ser Pro Ala Ser Ser Pro
70 75

Ser Ser Val Thr Tyr Pro Val Val Pro Gly Ser Val Asp
80 85 90

Glu Ser Pro Ser Gly Ala Leu Asn Ile Glu Cys Arg Ile
95 100

Cys Gly Asp Lys Ala Ser Gly Tyr His Tyr Gly Val His
105 110 115

Ala Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Thr Ile
120 125 130

Arg Leu Lys Leu Val Tyr Asp Lys Cys Asp Arg Ser Cys
135 140

Lys Ile Gln Lys Lys Asn Arg Asn Lys Cys Gln Tyr Cys
145 150 155

Arg Phe His Lys Cys Leu Ser Val Gly Met Ser His Asn
160 165

Ala Ile Arg Phe Gly Arg Met Pro Arg Ser Glu Lys Ala
170 175 180

Lys Leu Lys Ala Glu Ile Leu Thr Cys Glu His Asp Ile
185 190 195

Glu Asp Ser Glu Thr Ala Asp Leu Lys Ser Leu Ala Lys
200 205

Arg Ile Tyr Glu Ala Tyr Leu Lys Asn Phe Asn Met Asn
210 215 220

Lys Val Lys Ala Arg Val Ile Leu Ser Gly Lys Ala Ser
225 230

Asn Asn Pro Pro Phe Val Ile His Asp Met Glu Thr Leu
235 240 245

Cys Met Ala Glu Lys Thr Leu Val Ala Lys Leu Val Ala
250 255 260

Asn Gly Ile Gln Asn Lys Glu Ala Glu Val Arg Ile Phe
265 270

His Cys Cys Gln Cys Thr Ser Val Glu Thr Val Thr Glu
275 280 285

Leu Thr Glu Phe Ala Lys Ala Ile Pro Gly Phe Ala Asn
290 295

Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly
300 305 310

Val Tyr Glu Ala Ile Phe Ala Met Leu Ser Ser Val Met
315 320 325

Asn Lys Asp Gly Met Leu Val Ala Tyr Gly Asn Gly Phe
330 335

Ile Thr Arg Glu Phe Leu Lys Ser Leu Arg Lys Pro Phe
340 345 350

Cys Asp Ile Met Glu Pro Lys Phe Asp Phe Ala Met Lys
355 360

Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Ile Ser Leu
365 370 375

Phe Val Ala Ala Ile Ile Cys Cys Gly Asp Arg Pro Gly
380 385 390

Leu Leu Asn Val Gly His Ile Glu Lys Met Gln Glu Gly
395 400

Ile Val His Val Leu Arg Leu His Leu Gln Ser Asn His
405 410 415

Pro Asp Asp Ile Phe Leu Phe Pro Lys Leu Leu Gln Lys
420 425

Met Ala Asp Leu Arg Gln Leu Val Thr Glu His Ala Gln
430 435 440

Leu Val Gln Ile Ile Lys Lys Thr Glu Ser Asp Ala Ala
445 450 455

Leu His Pro Leu Leu Gln Glu Ile Tyr Arg Asp Met Tyr
460 465 468

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 468 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION : SEQ ID NO: 3:

Met Val Asp Thr Glu Ser Pro Ile Cys Pro Leu Ser Pro
5 10

Leu Glu Ala Asp Asp Leu Glu Ser Pro Leu Ser Glu Glu
15 20 25

Phe Leu Gln Glu Met Gly Asn Ile Gln Glu Ile Ser Gln
30 35

Ser Ile Gly Glu Glu Ser Ser Gly Ser Phe Gly Phe Ala
40 45 50

Asp Tyr Gln Tyr Leu Gly Ser Cys Pro Gly Ser Glu Gly
55 60 65

Ser Val Ile Thr Asp Thr Leu Ser Pro Arg Ser Ser Pro
70 75

Ser Ser Val Ser Cys Pro Val Ile Pro Ala Ser Thr Asp
80 85 90

Glu Ser Pro Gly Ser Ala Leu Asn Ile Glu Cys Arg Ile
95 100

Cys Gly Asp Lys Ala Ser Gly Tyr His Tyr Gly Val His
105 110 115

Ala Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Thr Ile
120 125 130

Arg Leu Lys Leu Val Tyr Asp Lys Cys Asp Arg Ser Cys
135 140

Lys Ile Gln Lys Lys Asn Arg Asn Lys Cys Gln Tyr Cys
145 150 155

Arg Phe His Lys Cys Leu Ser Val Gly Met Ser His Asn
160 165

Ala Ile Arg Phe Gly Arg Met Pro Arg Ser Glu Lys Ala
170 175 180

Lys Leu Lys Ala Glu Ile Leu Thr Cys Glu His Asp Leu
185 190 195

Lys Asp Ser Glu Thr Ala Asp Leu Lys Ser Leu Gly Lys
200 205

Arg Ile His Glu Ala Tyr Leu Lys Asn Phe Asn Met Asn
210 215 220

Lys Val Lys Ala Arg Val Ile Leu Ala Gly Lys Thr Ser
225 230

Asn Asn Pro Pro Phe Val Ile His Asp Met Glu Thr Leu
235 240 245

Cys Met Ala Glu Lys Thr Leu Val Ala Lys Met Val Ala
250 255 260

Asn Gly Val Glu Asp Lys Glu Ala Glu Val Arg Phe Phe
265 270

His Cys Cys Gln Cys Met Ser Val Glu Thr Val Thr Glu
275 280 285

Leu Thr Glu Phe Ala Lys Ala Ile Pro Gly Phe Ala Asn
290 295

Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly
300 305 310

Val Tyr Glu Ala Ile Phe Thr Met Leu Ser Ser Leu Met
315 320 325

Asn Lys Asp Gly Met Leu Ile Ala Tyr Gly Asn Gly Phe
330 335

Ile Thr Arg Glu Phe Leu Lys Asn Leu Arg Lys Pro Phe
340 345 350

Cys Asp Ile Met Glu Pro Lys Phe Asp Phe Ala Met Lys
355 360

Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Ile Ser Leu
365 370 375

Phe Val Ala Ala Ile Ile Cys Cys Gly Asp Arg Pro Gly
380 385 390

Leu Leu Asn Ile Gly Tyr Ile Glu Lys Leu Gln Glu Gly
395 400

Ile Val His Val Leu Lys Leu His Leu Gln Ser Asn His
405 410 415

Pro Asp Asp Thr Phe Leu Phe Pro Lys Leu Leu Gln Lys
420 425

Met Val Asp Leu Arg Gln Leu Val Thr Glu His Ala Gln
430 435 440

Leu Val Gln Val Ile Lys Lys Thr Glu Ser Asp Ala Ala
445 450 455

Leu His Pro Leu Leu Gln Glu Ile Tyr Arg Asp Met Tyr
460 465 468

What is claimed is:

1. Purified nucleic acid comprising the nucleotide sequence shown in SEQ ID NO. 1.

2. A vector comprising said nucleic acid of claim 1.

3. Recombinant PPAR expressed from said nucleic acid of claim 1.

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
ATGGTGGACA CGGAAAGCCC ACTCTGCCCC CTCTCCCCAC TCGAGGCCGG CGATCTAGAG AGCCCGTTAT CTGAAGAGTT CCTGCAAGAA ATGGGAAACA  100
M  V  D  T  E  S  P  L  C  P  L  S  P  L  E  A  G  D  L  E  S  P  L  S  E  E  F  L  Q  E  M  G  N  I

TCCAAGAGAT TTCGCAATCC ATCGGCGAGG ATAGTTCTGG AAGCTTTGGC TTTACGGAAT ACCAGTATTT AGGAAGCTGT CCTGGCTCAG ATGGCTCGGT  200
  Q  E  I  S  Q  S  I  G  E  D  S  S  G  S  F  G  F  T  E  Y  Q  Y  L  G  S  C  P  G  S  D  G  S  V

CATCACGGAC ACGCTTTCAC CAGCTTCGAG CCCCTCCTCG GTGACTTATC CTGTGGTCCC CGGCAGCGTG GACGAGTCTC CCAGTGGAGC ATTGAACATC  300
  I  T  D  T  L  S  P  A  S  S  P  S  S  V  T  Y  P  V  V  P  G  S  V  D  E  S  P  S  G  A  L  N  I

GAATGTAGAA TCTGCGGGGA CAAGGCCTCA GGCTATCATT ACGGAGTCCA CGCGTGTGAA GGCTGCAAGG GCTTCTTTCG GCGAACGATT CGACTCAAGC  400
E  C  R  I  C  G  D  K  A  S  G  Y  H  Y  G  V  H  A  C  E  G  C  K  G  F  F  R  R  T  I  R  L  K  L

TGGTGTATGA CAAGTGCGAC CGCAGCTGCA AGATCCAGAA AAAGAACAGA AACAAATGCC AGTATTGTCG ATTTCACAAG TGCCTTTCTG TCGGGATGTC  500
  V  Y  D  K  C  D  R  S  C  K  I  Q  K  K  N  R  N  K  C  Q  Y  C  R  F  H  K  C  L  S  V  G  M  S

ACACAACGCG ATTCGTTTTG GACGAATGCC AAGATCTGAG AAAGCAAAAC TGAAAGCAGA AATTCTTACC TGTGAACATG ACATAGAAGA TTCTGAAACT  600
  H  N  A  I  R  F  G  R  M  P  R  S  E  K  A  K  L  K  A  E  I  L  T  C  E  H  D  I  E  D  S  E  T

GCAGATCTCA AATCTCTGGC CAAGAGAATC TACGAGGCCT ACTTGAAGAA CTTCAACATG AACAAGGTCA AAGCCCGGGT CATCCTCTCA GGAAAGGCCA  700
A  D  L  K  S  L  A  K  R  I  Y  E  A  Y  L  K  N  F  N  M  N  K  V  K  A  R  V  I  L  S  G  K  A  S

GTAACAATCC ACCTTTTGTC ATACATGATA TGGAGACACT GTGTATGGCT GAGAAGACGC TGGTGGCCAA GCTGGTGGCC AATGGCATCC AGAACAAGGA  800
  N  N  P  P  F  V  I  H  D  M  E  T  L  C  M  A  E  K  T  L  V  A  K  L  V  A  N  G  I  Q  N  K  E

GGCGGAGGTC CGCATCTTTC ACTGCTGCCA GTGCACGTCA GTGGAGACCG TCACGGAGCT CACGGAATTC GCCAAGGCCA TCCCAGGCTT CGCAAACTTG  900
A  E  V  R  I  F  H  C  C  Q  C  T  S  V  E  T  V  T  E  L  T  E  F  A  K  A  I  P  G  F  A  N  L

GACCTGAACG ATCAAGTGAC ATTGCTAAAA TACGGAGTTT ATGAGGCCAT ATTCGCCATG CTGTCTTCTG TGATGAACAA AGACGGGATG CTGGTAGCGT 1000
D  L  N  D  Q  V  T  L  L  K  Y  G  V  Y  E  A  I  F  A  M  L  S  S  V  M  N  K  D  G  M  L  V  A  Y

ATGGAAATGG GTTTATAACT CGTGAATTCC TAAAAAGCCT AAGGAAACCG TTCTGTGATA TCATGGAACC CAAGTTTGAT TTTGCCATGA AGTTCAATGC 1100
  G  N  G  F  I  T  R  E  F  L  K  S  L  R  K  P  F  C  D  I  M  E  P  K  F  D  F  A  M  K  F  N  A

ACTGGAACTG GATGACAGTG ATATCTCCCT TTTTGTGGCT GCTATCATTT GCTGTGGAGA TCGTCCTGGC CTTCTAAACG TAGGACACAT TGAAAAAATG 1200
L  E  L  D  D  S  D  I  S  L  F  V  A  A  I  I  C  C  G  D  R  P  G  L  L  N  V  G  H  I  E  K  M

CAGGAGGGTA TTGTACATGT GCTCAGACTC CACCTGCAGA GCAACCACCC GGACGATATC TTTCTCTTCC CAAAACTTCT TCAAAAAATG GCAGACCTCC 1300
Q  E  G  I  V  H  V  L  R  L  H  L  Q  S  N  H  P  D  D  I  F  L  F  P  K  L  L  Q  K  M  A  D  L  R

GGCAGCTGGT GACGGAGCAT GCGCAGCTGG TGCAGATCAT CAAGAAGACG GAGTCGGATG CTGCGCTGCA CCCGCTACTG CAGGAGATCT ACAGGGACAT 1400
  Q  L  V  T  E  H  A  Q  L  V  Q  I  I  K  K  T  E  S  D  A  A  L  H  P  L  L  Q  E  I  Y  R  D  M

GTACTGA                                                                                                      1407
Y  X
```

FIG. 1

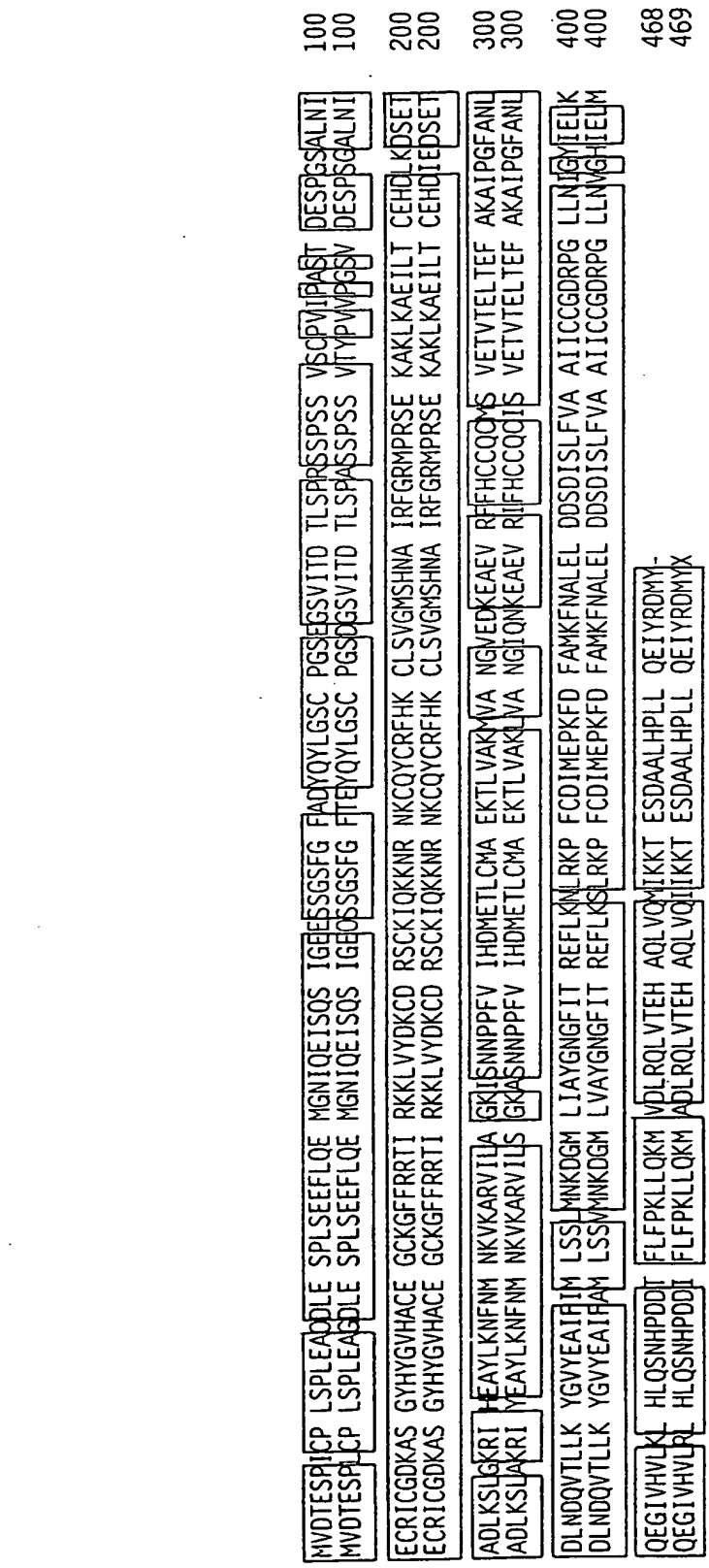

MVDTESPLCP LSPLEADDLE SPLSEEFLQE MGNIQEISQS IGEESSGSFG FADVQYLGSC PGSEESGSVITD TLSPRSSPSS VSCPVIPAGT DESPSSALNI 100
MVDTESPLCP LSPLEADDLE SPLSEEFLQE MGNIQEISQS IGEQSSGSFG FTEVQYLGSC PGSDQSGSVITD TLSPASSPSS VTYPVVPPGV DESPSGALNI 100
ECRICGDKAS GYHYGVHACE GCKGFFRRTI RKKLVYDKCD RSCKIQKKNR NKCQYCRFHK CLSVGMSHNA IRFGRMPRSE KAKLKAEILT CEHDLKDSET 200
ECRICGDKAS GYHYGVHACE GCKGFFRRTI RKKLVYDKCD RSCKIQKKNR NKCQYCRFHK CLSVGMSHNA IRFGRMPRSE KAKLKAEILT CEHDIEDSET 200
ADLKSLGKRI HEAYLKNFNM NKVKARVILA GKISNNPPFV IHDMETLCMA EKTLVAKMVA NGVEDKEAEV RFFHCCQQVS VETVTELTEF AKAIPGFANL 300
ADLKSLAKRI YEAYLKNFNM NKVKARVILS GKASNNPPFV IHDMETLCMA EKTLVAKLVA NGIQNKEAEV RIFHCCQQIS VETVTELTEF AKAIPGFANL 300
DLNDQVTLLK YGVYEAIFIM LSSLMNKDGM LIAYGNGFIT REFLKNLRKP FCDIMEPKFD FAMKFNALEL DDSDISLFVA AIICCGDRPG LLNIGYIEIK 400
DLNDQVTLLK YGVYEAIFAM LSSMMNKDGM LVAYGNGFIT REFLKSLRKP FCDIMEPKFD FAMKFNALEL DDSDISLFVA AIICCGDRPG LLNVGHIEIM 400
QEGIVHVLKL HLQSNHPDDI FLFPKLLQKM VDLRQLVTEH AQLVQMIKKT ESDAALHPLL QEIYRDMY- 468
QEGIVHVLRL HLQSNHPDDI FLFPKLLQKM ADLRQLVTEH AQLVQIIKKT ESDAALHPLL QEIYRDMYX 469

FIG. 2